

Watcha say?- developing and implementing standards for assessing speaking

Dr. Craig L. Dicker

U.S. Embassy, Ankara

DickerCL@State.gov

<http://turkey.usembassy.gov>

The “dreaded” CEF

- The CEF is a straight-jacket that will force all institutions to conform to a single set of standards
- The CEF is NOT sensitive to the different backgrounds, needs and educational systems found in different societies
- The CEF is becoming the only criteria to measure test validity

What the CEF is NOT....

- A system that reflects an order or hierarchy of language acquisition that has empirical support
- A proficiency/standards based curriculum that guides teachers and learners with regard to what should be taught and when
- A set of proficiency descriptors that can be used as a rating scale to assess a person's language proficiency level
- A framework upon which existing language exams can be mapped so that a score on one exam (e.g. the TOEFL) can be equated with a CEF score

Intended use and objectives of the CEF

The CEF aims to provide a common point of reference- not an assessment tool. The CEF is meant to be comprehensive but users must be selective in using it to...

- plan for language certification programs (syllabus content, assessment criteria)
- plan for language learning programs (objectives, content, placement)
- plan for self-directed learning (materials, assessment, objectives)

The CEF can help...

- Move language education programs away from a “knowledge of the target language” focus to a “ability to use the target language” focus
- Empower programs in language education to develop skills-oriented curricula, exams and instructional materials
- Create a system of continuous assessment where class participation, self and peer assessment and completion of course requirements count much more than successfully passing a single exam.

The CEF can also help to...

- Encourage teachers to use more interactive, “communicative” methods in their language classes.
- Increase inter-institutional and international student mobility by providing a common framework allowing users to assess the similarities and differences among systems and give meaning to courses completed.
- Address workplace language needs by producing graduates with the language skills needed in global markets.

Competence vs. Performance

- Definition of "Competence"-
What a person can actually do
- Can NOT be directly measured
- A purely psychological construct- no reference to the social nor environmental variables
- Definition of "Performance"- what a person actually does in a particular setting (e.g. taking a test)
- Competence can only be measured via one's performances
- One's performance can never equal one's ideal competence- the former will always fall somewhat short

Responsibilities of Test Givers

- To maximize the test takers' performances so that they come as close to their real competence as possible
- To vary tasks in an effort to find those tasks they feel most comfortable with
- To ensure that the content of the test (issues/subjects dealt with) do not penalize test takers who are not intimate with those subjects/issues
- To focus on the language production of the test takers- not to be influenced or swayed by their personalities

Validity and Reliability

- Validity: Tests assess what they are designed to assess (e.g. oral language skills and NOT subject area knowledge)
- Reliability
 - Inter-rater: Language production (data) is scored the same by DIFFERENT raters
 - Intra-rater: Language production (data) is scored the same by the SAME rater but at different times

Oral Test Task Types

- Narrative (e.g. story telling based on pictures, graph explanations, etc.)
- Interactive (e.g. Proficiency Interviews)
- Information Gap (e.g. describing a picture to another who draws it)
- Role Play (e.g. at the Travel Agency)
- Oral Translation-Interpretation

Curriculum vs Standards-Referenced Programs

- Curriculum-based programs...
 - consist of a series of related items to “cover”- often dictated by a textbook-curriculum
 - can be grammatical, functional, thematic, task-based, procedural, situational, etc. or a combination of the above
- Standards-referenced programs...
 - revolve around a set of articulated “behavioral objectives” which the course is designed to help learners achieve
 - are often institution-specific and reflect the institution’s stated goals and assessed learner needs

Curriculum vs Standards-Referenced Exams

- Curriculum-based exams are designed to assess the learners' mastery of the content of the curriculum. The nature of the exam should reflect the content of the curriculum
- Standards-referenced exams are designed to assess the learners' mastery of the behavioral objectives set out for the course. These objectives are based on the institution's standards for particular levels/skills

CEF Overall Oral Production Standards

| | |
|---------------------------------------|---|
| C2 Mastery | Can produce clear, smoothly flowing well-structured speech with an effective logical structure |
| C1 Effective Op.Proficiency | Can give clear, detailed descriptions and presentations on complex subjects, integrating sub-themes, developing particular points and ending w/ a conclusion |
| B2+ Vantage+ | Can give clear, systematically developed descriptions and presentations, with appropriate highlighting of significant points and relevant supporting detail |
| B2 Vantage | Can give clear, detailed descriptions and presentations on a wide range of subjects related to field of interest, expanding and supporting ideas w/relevant examples |
| B1 Threshold | Can reasonably fluently sustain a straightforward description of one of a variety of subjects within field of interest, presenting it as a linear sequence of points |
| A2 Waystage | Can give a simple description or presentation of people, living-working conditions, daily routines, likes/dislikes as a short series of simple phrases and sentences |
| A1 Breakthrough | Can produce simple mainly isolated phrases about people and places |

Comparing and Contrasting B1 and B2 Oral Standards

| | B1 | B2 |
|-----------------|---|--|
| Range | Can get by with sufficient vocabulary to express him/herself with circumlocutions on topics like family, hobbies and interests, work, travel and current events | Has sufficient range of language to be able to give clear descriptions, express viewpoints on most general topics without much conspicuous searching for words, using some complex sentence forms to do so |
| Accuracy | Uses reasonably accurately a repertoire of frequently used routines and patterns associated with more predictable situations | Shows a relatively high degree of grammatical control. Does not make errors which cause misunderstanding, and can correct most of his/her mistakes |

Council of Europe (2001). Common European Framework of Reference for Languages Cambridge University Press

Comparing and Contrasting B1 and B2 Oral Standards

| | B1 | B2 |
|--------------------|--|--|
| Fluency | Can keep going comprehensibly, even though pausing for grammatical and lexical planning and repair is very evident | Can produce stretches of language with a fairly even tempo; although can be hesitant as he/she searches for patterns/expressions. |
| Interaction | Can initiate, maintain and close simple face-face conversations on topics that are familiar or of personal interest. Can repeat back part of what has been said to confirm understanding | Can introduce discourse, take appropriate turns and end conversations. Can help the discussion along on familiar ground confirming comprehension, inviting others in, etc. |
| Coherence | Can link a series of shorter, discrete simple elements into a connected, linear sequence of points | Can use a limited # of cohesive devices to link utterances into clear, coherent discourse |

Oral standards generally take the following into consideration

- nature of vocabulary used (range)
- pronunciation and comprehensibility
- syntactic complexity and accuracy
- fluency
- types of situations handled
- coherence

Aural standards take into consideration...

- subjects/topics (from basic survival to abstract professional)
- source of input (e.g: face-face foreigner talk, mass media, academic lectures)
- length of input (e.g. isolated words, sentences, extended discourse)
- syntactic complexity
- vocabulary usage

Norming

- Objective- to ensure that at least 90% of all raters give same score to a given interview (inter-rater reliability)
- Procedure- play audiotapes of interviews where interlocutors are unknown to teachers
- Review ratings, discuss and continue- fine tuning of criteria descriptions for each level can occur during norming sessions
- Exam checks carried out by roaming monitors and spot audiotaped interview inspections

Interview training for oral assessment

- Experienced interviewer demonstrates phases of interview
- Teachers put into groups
- Learners from different levels sent to each group
- 15-20 minute interviews conducted (audiotaped), followed by 10-15 minute small group discussion of interview style
- Full group meeting to discuss difficulties

Interview Phases

- Warm Up- put students at ease
- Level Check- determines the level the learner can comfortably function in - establishes base line
- Probes- requires learner to try to function at a level beyond the established baseline- if successful, a new, higher baseline is established. Probes continue till a “ceiling” (point when language breaks down) is established
- Wind Down- return to baseline level and end interview

Sample Interview Questions

- Life-Personal (where do you live?)
- Life-Personal Concrete-Descriptive (Describe a wonderful place you have visited)
- Life Personal Abstract (What characteristics do you look for in a friend?)
- Non-Personal Concrete (What qualifications are necessary to practice medicine?)
- Non-Personal Abstract (Why do you think socialism failed in Eastern Europe?)

USEFUL REFERENCES

- Common European Framework of Reference for Languages:
http://www.coe.int/T/E/Cultural_Cooperation/education/languages/Language_Policy
- Fulcher, G. (2003). Testing Second Language Speaking. London: Longman/Pearson Educational
- Davidson, F. and Lynch, B. (2002). Testcraft: A Teacher's Guide to Writing and Using Language Test Specifications. New Haven: Yale University Press.